Eilam Gross, Ofer Vitells
10 December, 2010

Banff challenge 2a results

1 Problem 1

The problem consists of a gaussian-distributed signal events with unknown rate and location on top of a background that follows an exponential distribution. The background rate comes with an uncertainty of 10%. The statistical uncertainty of determining the background rate from the data is expected however to be considerably smaller so we ignore this information and treat the background rate as a completely unknown parameter.

We analyze the data using the profile likelihood ratio approach, following [1]. We use an unbinned likelihood defined as:

$$\mathcal{L}(N_s, N_b, M) = \prod_i \frac{N_s f_s(x_i; M) + N_b f_b(x_i)}{N_b + N_s} \times Poiss(N|N_s + N_b) \tag{1}$$

where N_s, N_b are the total expected number of signal and background events, respectively, and M is the signal location. f_s, f_b are the (normalized) signal and background pdf's.

The test statistic for discovery is given by

$$q_0 = \begin{cases} -2\log\lambda(0) & \hat{N}_s > 0 \\ 0 & \hat{N}_s < 0 \end{cases} \tag{2}$$

where

$$\log\lambda(N_s) = \log\frac{\mathcal{L}(N_s, \hat{\hat{N}}_b, \hat{\hat{M}})}{\mathcal{L}(\hat{N}_s, \hat{N}_b, \hat{M})} \tag{3}$$

To account for the 'look elsewhere effect', we use the procedure described in [2]. We calculate q_0 as a function of M and estimate the number of upcrossings at a level $q_0 = 0.5$ from a set of 100 randomly generated background datasets, from which we estimate this number to be 2.28 ± 0.1 . The p-value at a higher level is then estimated by:

$$P(q_0 > c) = \frac{1}{2}P(\chi^2 > c) + 2.28e^{-(c-0.5)/2} \tag{4}$$

The critical value of q_0 is derived from the above formula by requiring that the p-value is 0.01, and we obtain $q^c = 11.43$. Thus an evidence for the signal is claimed if $q_0 \geq 11.43$.

1.1 Sensitivity

The power of the test is estimated by MC simulation of datasets containing a signal. We find that the power for the three cases $(D, E) = (1010, 0.1), (137, 0.5), (18, 0.9)$ is 35%, 46%, and 19%, respectively.

1.2 Confidence intervals

Confidence intervals for the signal location and rate are derived by requiring that $\Delta 2 \log \lambda = 1$, when the relevant parameter (N_s , M) is shifted. For the signal rate, this approximation is expected to break when the signal is small due to background fluctuations far from the true signal location. To guarantee coverage for small signals we therefore additionally set the lower bound on the signal rate to be zero when $P(q_0 \leq q_0^{observed}|H_0) = 68\%$.
(N_s is related to the rate parameter D defined in the problem via $D = \frac{N_s}{\sqrt{2\pi}\sigma}$.)

2 Problem 2

The analysis of this problem follows generally that of problem 1, except that there is no 'look elsewhere effect'. We use a binned likelihood function:

$$\mathcal{L}(N_s, N_b, M) = \prod_i Poiss(n_i|N_s s_i + N_b^1 b_i^1 + N_b^2 b_i^2) \tag{5}$$

where s_i, b_i^1, b_i^2 are the expected fractions of signal and background events in the i-th bin (this is taken from the provided MC simulations).

The test statistic is defined similarly to problem 1. Here the number of events is large enough so everything is calculated using asymptotic properties. The critical value of q_0 that corresponds to a p-value of 0.01 is found from the χ^2 cdf, which gives $q^c = 5.41$. Thus an evidence for the signal is claimed if $q_0 \geq 5.41$.

2.1 Sensitivity

The sensitivity is estimated from the 'Asimov' dataset as defined in [1], and assuming the asymptotic non-central χ^2 distribution for q_0 under H_1. That is, the power is given by: Power=$1 - F_{\chi^{2\prime}}(q^c; q^A)$, where $F_{\chi^{2\prime}}(x; \Lambda)$ is the cumulative distribution of a non-central χ^2 with non-centrality parameter Λ , and q^A is the Asimov value of q_0. For a signal rate of $D = 75$ this gives a power of 81.5%.

2.2 Confidence intervals

The confidence interval for the signal rate are estimated by requiring that $\Delta 2 \log \lambda = 1$, when N_s is shifted. (N_s is restricted to be positive, so the interval bounds are always positive)

References

[1] Glen Cowan, Kyle Cranmer, Eilam Gross, Ofer Vitells, *Using the Profile Likelihood in Searches for New Physics*, [arXiv:1007.1727v2].

[2] Eilam Gross, Ofer Vitells, *Trial factors for the look elsewhere effect in high energy physics*, Eur. Phys. J. C, **70**, 1-2, (2010).